Suggested Structure and Structuring of an Empirical Master's Thesis in Finance

by Lammertjan Dam

The first part of this teaching note is meant to outline the generic structure of a Master's thesis in Finance. I would like to point out that this is merely a guideline, any deviations from the suggested schedule are possible. Second, in practice the sections of a thesis are never actually written down in the order in which they are presented. In the second part of this note I will give an indication of how researcher usually progress when writing down their research (or at least how I usually do it).

Part 1: What should I write down?

1. Introduction (2-3 pages)

The introduction should prepare the reader for the main body of the thesis. The key of good writing is to make sure that the reader is actually going to read what he *expects* to read. A constant confirmation of the reader's expectations of what is going to come makes the thesis a pleasent read; it gives the reader the feeling that he understands what is going on, and that is the goal of communicating your work. This does not only hold for the introduction but for the thesis as a whole.

Questions that should be answered in the introduction are:
- What is the topic of your research?
- What is it specifically that you add to this area of research?
- What is the question you are trying to answer?
- Why is this interesting?
- How do you do it and what are the specifics of the data that you use? (country/ year?)

2. Literature review/ Background (5-10 pages)

The literature review should not be a simple list of other articles and/or books that have also dealt with the same topic. Instead, try to answer your research question by using the existing literature. So before you actually start

answering your research question with data, you want to develop some idea of what you could expect to find based on previous research.

Question that should be answered are:
- What are the (seminal) theories in the existing literature?
- What do these theories predict (in the context of your research)?
- Are there important debates in the literature?
- What empirical evidence is found in line with these theories?
- What are other important factors that should be considered (control variables!)?
In sum, based on theoretical and empirical work, what are you likely to find?

3. Methodology and/or Hypothesis (2-5 pages)

If you have written a decent literature review the hypothesis follows naturally. However, in this section you should be even more specific about how you are actually going to test this hypothesis.

Question that should be answered are:

- What type of data do you need? (Note: you do not have to describe the data in full detail yet!)
- What statistical relationship are you going to test using this data? (give equation).
- What are important control variables and what do you expect to find for these variables? (In the case of regression analysis).
- How does the significance and the sign of the various parameters relate to the more general hypothesis that you want to test?
- What justifies this approach?

In fact the goal is to formalize the research question into a statistical hypothesis.

4. Data and descriptive statistics (4-8 pages)

In this section you list your data sources and try to give the reader some feeling for the data you have collected. Remember, you have played around with the data a lot yourself, but the reader only has your thesis. Therefore, you do not only tell the reader what your data sources are, but you should also give some descriptive statistics, such as:

- Summary statistics (for both dependent and independent variables), i.e. Mean/St.Dev/Median/Min/Max/Number of Observations
- Summary statistics of various subsamples, e.g. by country or year.
- You could also split the sample according to a threshold measure of some independent variables to see in which direction the data points regarding the research question.

What statistics you are going to present in the main text and what you put in an appendix is an important choice that you have to make. But always keep in mind that you want to inform the reader as well as possible, and you also don't want to confuse him, so really think about the selection of tables you want to present. Tables with statistics that are beside the main point of your thesis should be put in the appendix. For example, a test for normality on your residuals in an OLS regression is often best put in the appendix. Also, if you link various datasets, it is very likely that you have to drop observations because of missing data. Report this data selection procedure as well.

Finally, repeat in the main text what you present in the tables. This kind of seems like "double work", but it is important. This way, you confirm what the reader thinks he is seeing in the table. You do not have to report the complete table entrywise, but think of statements such as *"The average daily stock return, reported in column A of table 1, varies between -0.2% and 1.3% for the various stock markets, and there is also quite some variation in stock returns, as can be seen from the standard deviation, reported in column B, which is around 2% for each stock market".*

5. Results (4-8 pages)

Here you present your main results. Try to present your results using a maximum of 2 or three *main* tables; any other statistical analysis should be either labeled as a robustness check (for example, various sample splits such as different periods or different subsets of countries) or other resulst that support the main point (for example, other statistical relationships that are predicted by the same theory).

Of course, the same guidelines for reporting the tables as described in the *data section* hold. In addition however, it is even more important that you give economical meaning to the results. Do not simply say: *"the sign of beta is positive and significant"*, but *"the sign of beta is positive and significant, which indicates that investors that face a systematic risk ares rewarded with a higher stock return."* Essentially, you are slowly becoming less formal again and you try to

place your results in the context of the more general research area you started with in the introduction. Also, if you have unexpected findings, can you explain it?

A note on presenting tables

First of all, I feel that tables should be *self-contained.* That is, someone should be able to understand everything that is presented in a table without having read your thesis at all. Tables draw people's attention, so they should be attractive. Of course you explain everything in the text as well, but this is only to confirm that the reader really has understood the table. Also think about the layout. Do you really need 10 digits? Also, do not forget about the units.

To illustrate my point, consider the following table:

CR	P	IRR
NL000233.5%	223.4563423	0.03323478677
NL000233.7%	213.2323423	0.03234234327
NL000233.8%	223.4563423	0.03456534535
NL000233.9%	313.2343423	0.03435345599
NL000233.1%	115.4563423	0.03344534599
NL000233.4%	334.5646899	0.03576563499

And compare it to the table below, which contains the same information.

Table 1. Dutch state bonds and yields

Coupon rate (%)	Price (€)	Internal Rate of Return* (%)
3.5	*223*	*3.32*
3.7	*213*	*3.23*
3.8	*223*	*3.46*
3.9	*313*	*3.44*
3.1	*115*	*3.34*
3.4	*334*	*3.58*

* Note: The Internal Rate of Return (IRR) is calculated using $P= c/(1+y) +\ldots +c/(1+y)^n$, where P is price, c is the coupon rate and y is equal to IRR.

Which table do *you* find more clear?

6. Conclusion (2-3 pages)

First, repeat what you have done and what your results are. Second, describe what your results mean in the context of the existing literature. Are your results in line with the literature? Do they support theory? Third, list limitations of your thesis (data problems? methodological issues?). Finally, list some directions for future research.

7. Appendix (2-15 pages; it really depends)

Here you can put variable definitions, additional descriptive statistics, other robustness checks, pre- or post-estimation diagnostics, mathematical derivations, statistical theory, etc.. Anything which is beside the main point, but may be useful for the interested reader.

Part 2: In what order should I write things down?

I think the most efficient way for a master student to produce his thesis is the following:

1. find a subject of interest. Find some related literature (anywhere between 3-20 papers) and make an overview, using a table for example, of what others have done and what they have found.

2. Select an empirical paper which you like the most and understand well. Use this methodology and adapt it in such a way that it will fit your research question.

3. By now you have an idea of the equation you are going to estimate, so figure out what data you need to collect. Also, based on the literature you have read, you should be able to list the most important control variables.

4. After you have collected the final dataset, make many tables with summary statistics. The tables with descriptive statistics should be the first thing that you really write down "neatly". Since you are going to use the data – no matter what the results turn out to be – these tables will always be useful, even if you choose

to put some in an appendix later. At this point, your dataset will not change anymore, so the description of your data will also not change. It means that you can start writing part of the *data and descriptive statistics* section already.

5. After you have done the actual analysis, also make nice tables for your main results. You can also start writing part of the *results* section.

6. Finalize the *data and descriptives* and the *results* section, and at the same time start writing down the methodology section. Usually, the research question has slightly changed by now, so do another thourough literature review and start writing that section as well, but this time you do not approach it as a list, but you are really trying to write an interesting story. Check how other authors do this in terms of style and presentation.

7. Finally write the conclusion and the introduction.

8. Let others read your thesis and re-read it yourself after you have put it away for a while. Especially the introduction and the conclusion should be very clear: most people read this (first).

9. Do not underestimate the importance of layout and presentation. Give tables clear titles (Not: "results, but be specific, e.g. "Risk and return for stocks listed at the AEX stock index".) Do not start a new paragraph (dutch: alinea) to often. Give tables, pages, and figures numbers. Spell-check your document.

Final remarks:

- Although the order of putting the thesis together starts with the data, it does not mean that your research question should be data-driven.
- Everything in your thesis should be as clear as possible. Do not be afraid to state the obvious. It is a good thing if your readers think:"yeah, yeah, yeah, this is easy stuff." It means your thesis is well-written.